

AD710933

11  
①  
Presented to:  
Office of Naval Research (Code 453)  
Department of the Navy  
Washington 25, D. C.

## DECISION MAKING

C. L. Winder

Technical Report Number I  
Office of Naval Research Project NR 150-087,  
Contract Nonr 225(01)

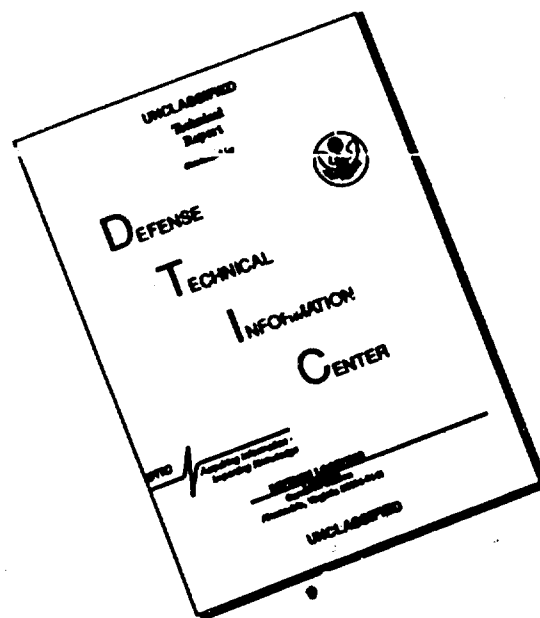
Department of Psychology  
Stanford University  
Stanford, California  
March, 1953

SEP 8 1970

This document has been approved for public release and sale;  
its distribution is unlimited.

Reproduced by the  
CLEARINGHOUSE  
for Federal Scientific & Technical  
Information Springfield Va 22151

# DISCLAIMER NOTICE



**THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.**

### Preface

This study was conducted with the assistance of John Vitale and Wirt Wolff. Helene Veltfort and Joe Fortier were also associated with the project.

## Chapter I - Introduction

This technical report is an account of some of the investigations conducted under terms of Project NRL50-087, Contract Nonr 225(01) between the Office of Naval Research and Stanford University. The general goal of the project was to study ~~certain~~ relationships between decision making and personality characteristics. It was recognized explicitly that the project would be largely exploratory because the subject of concern is a relatively unexplored area.

The context out of which this study arose was the desire to understand better the factors determining the individual differences in formulations and executions of policies or strategies for dealing with problem situations. The word "problem" is used here in a broad sense; i.e., we do not refer to problems of a specific mathematical or logical kind, but to situations where a reasoned approach would presumably be more desirable than an unconsidered or impulsive approach. We refer to adaptation in situations in which the context--present, past, and future--is taken into account, and where the reaction is not an automatic one. This is not to say that unconscious, or emotional, or any other class of factors is considered unimportant in contributing to the determination of these reactions. Clearly, this is a major aspect of human behavior and some narrowing of the area of investigation was necessary.

One characteristic of general problem situations is that many different possibilities for response exist, and only some fraction of the possibilities can be utilized. In a problem situation, the person alters himself and the situation by selective behavior. The extent of variation of behavior possible in problem situations is a function of both the external situation and the particular person involved. (Situations

involving more than one person are subsumed by considering other people as part of the situation of any single individual.)

If we accept this general view, a consideration of primary moment is the explanation of the choice of that behavior to be engaged in out of those behaviors possible. Hypothetically, such a choice can be: more or less appropriate, quickly or slowly made, more or less inclusive in terms of time and elements, related by reason in various degrees to antecedent and subsequent choices, etc.

Our attention is centered, then, on decisions. Involved in a decision is a functional awareness of the possibility of actions or inaction, the formulation of alternatives, the selection of an alternative, and the actualization to some extent of that alternative. There then follows the functional awareness of some change. (By using the concept "functional awareness" we temporarily bypass questions of the degree of consciousness and the degree of overtness in the decision making process and consequent behavior. This statement of the above steps in a decision is not meant to be exhaustive nor is it a theory, but is simply an attempt to make explicit the general approach followed in setting out to study decision making at an empirical level.)

One area of investigation which is relevant is the relationship between adequate behavior in a problem solving situation and intelligence. This area has not been thoroughly explored, but sufficiently so that we know the relationship is positive but not exhaustive in terms of the proportion of variance explained. We would expect decision capacity and ability to be explained in part by intellectual capacity. Yet it is a common conviction among those concerned with this area that more is involved than intellectual capacity as now measured. To state the point in an alternative way, among a group of persons who score high on intelligence tests, i.e. are similar in intellectual capacity, there are noticeable

differences in the adequacy of decisions made. It seems a justifiable assertion to say that there is much about decision making which is not understandable in terms of its relationship to measured intelligence.

It should be clear that the term "decision making" as used here is a descriptive concept, not necessarily a denotative one and that decision making refers to an aspect of the integrated function of the organism; hence, focusing on this aspect will necessarily involve some fictions, as does any analytic process.

If individual differences in decision making are not adequately explained by the intellectual capacity of the person, what other factors are involved? The general area of personality characteristics suggests itself. There is no absolutely clear line of demarcation between measures of intellectual capacity and measures of personality characteristics. Nevertheless, the distinction is tenable and useful. Common sense observation suggests the importance of personality characteristics in connection with decision making. Some people are actually characterized as decisive or indecisive; or, a person may be spoken of as capable of taking administrative responsibility well; or, many women are said to be unable to make up their minds; or, a person may be said to be consistent in what he chooses while another is described as making many unrelated or incongruous choices.

The decision making behavior of our fellow men is recognized to be variable, and is itself viewed in some cases as a personality characteristic. We might almost have entitled the project "Studies of Decision Making as Related to Other Personality Characteristics."

The available measures of personality characteristics are not as fully developed as would be desirable. Therefore, some effort must be devoted to clarifying and evaluating certain measures of personality

which seem relevant to the problem at hand. Wherever good personality measures already exist which might prove appropriate, these are employed.

Suggestive of the need to explore the variety of decision making functions is the fact that a distinction is frequently made between choice and judgment. Also, some decisions seem to be made primarily in terms of external goals while others seem to be guided primarily by internal goals. Other common sense distinctions could be mentioned in this context, such as the possibility that some individuals characteristically make decisions piecemeal while others make decisions involving more complete courses of action. The development of a variety of measures of decision making and an analysis of the relationships among these would empirically establish some types or facets of decision making and is accordingly an aim of the project. On such an analysis of the problem area, further research possibilities will rest.

In summary, this is frankly an exploratory project which has as an overall goal the fuller understanding of decision making. In order to accomplish this aim, we must be concerned with the development of measures of decision making, with efforts being made to specify types of decision making, and to clarify and develop pertinent personality measures. Also we will be concerned with how decision making capacity and ability is related to overall personality and to other psychological capacities.

Succeeding chapters of this report will contain accounts of studies undertaken on the basis of this general viewpoint.

## Chapter II

### Individual Differences in Inconsequential Judgment Making

As was indicated in the first chapter, decision making is viewed as including judgment function(s). Our investigation of judgment was approached as an individual differences problem for two reasons: first, the general framework of the present project is that of the psychology of the individual, as an integrated, functioning organism; second, there is a considerable fund of knowledge regarding the methodology and phenomena of judgment making from a psychophysical point of view which furnishes an adequate background for such an individual differences study.

An inconsequential judgment situation was selected for study on grounds analogous to those for the use of "projective" techniques. The individual, faced with instruction to make a judgment remote from any real-life situation should reveal any stable, basic judgment-making characteristics he might possess. The effort was to test the assumption that the usual, intelligent young adult human has as part of his psychological equipment some characteristic patterns of judgment response or decision making propensities (the question of the contribution of learning in determining those patterns is not the question under investigation at the moment.)

#### Method

Each subject participated in this judgment experiment, as well as a variety of personality assessment situations. Only the judgment experiment will be reported in this chapter.

Two judgment problems were set for the subjects. In one, judgment of the brightness of a tone was asked for and in the other the problem was a judgment of the density of the same tone. (These problems will be

referred to subsequently as "Psychophysical Tasks.") The two experiments were conducted one week apart, each being accompanied by other procedures to be discussed in the succeeding chapters. There is no reason to believe that the accompanying procedures affected the judgment behavior in any systematic way; however, this possibility was not evaluated. Subjects were permitted a few minutes to adapt to the dimly lighted experimental room. The same dim illumination was used for both the judgment tasks.

The well-known psychophysical method of limits was employed for the psychophysical tasks, each subject making 8 series of judgments or runs in each judgment problem.

The tone used was a relatively pure tone of about 500 cycles per second, produced by a standard laboratory oscillator. The tone was monitored by means of an oscilloscope to insure reasonable uniformity. This tone was kept within acceptable limits of variation in intensity and frequency, but auditory laboratory standards of control were not necessary since audition was not the subject matter of the study. Thus, our general conditions could be reproduced in almost any psychological laboratory. Control of external distractions was made possible as the subjects were studied in a sound-resistant room which excludes all ordinary sounds. This arrangement also protected the environment from undue disturbance.

In the judgment of brightness problem, the subject was presented with the constant tone, and told that he would then be presented with a series of 16 illuminations of a ground glass screen, he being asked to say of each illumination that its brightness was greater, equal, or less than that of the tone. Pre-tests were conducted to find a range of illumination which would permit almost all subjects to make this judgment, perceiving some illuminations in the series as brighter and some as less

bright than the tone. Each illumination was sufficiently discriminable from every other illumination so that every subject gave judgments with no reversals.\*

In the method of limits, the series of matching stimuli (illuminations in this case) are presented, one at a time and in order of magnitude, starting with the one of largest magnitude on alternate runs and with the stimulus of least magnitude on the other runs.

The judgment of density experiment was conducted in the same manner. The matching stimuli were brass cylinders of exactly the same size and appearance but of different weights. The series chosen consisted of 16 weights which were selected so that almost all subjects would be able to make the required equation, but in contrast to the illumination stimuli, these weights were selected so that there could be some confusion as to which of any two successive weights were heavier.

The subjects were clearly told in each judgment situation that the stimuli would be presented in series of increasing or decreasing magnitude. The subjects were also instructed, in each experiment, to make every judgment separate from every other judgment; i.e., to compare the brightness or density of each separate stimulus with that of the tone.

These judgment tasks were rather stressful for the subjects who, nevertheless, could be kept to the task.

#### Subjects

All (N=45) subjects were students at Stanford University, most of whom were enrolled in the first course in psychology. Each student in this course is asked to participate in a number of experiments to appreciate

---

\* A reversal is an instance where an illumination which is actually of greater physical intensity is perceived as being less bright than one of less physical intensity, or the converse.

the complexities and subtleties of psychological experimentation. Such subjects are usually motivated to cooperate in the procedures wholeheartedly. When students had already participated in other experiments and did not need to accumulate all of the hours required in this project (5 hours total), they were paid a nominal fee for participating. This arrangement met with enthusiasm and promoted cooperation on the part of the subjects.

Both men and women were used as subjects, and no age restrictions were imposed. Subjects averaged around 19 years of age.

#### Results

Judgments of Brightness: Two scores were utilized here, namely: the Time Score ( $T_1$ ) and the Interval of Uncertainty Score ( $I U_1$ ).

$T_1$  for each subject is the average time per series, i.e., the time from the initiation of a series or run of stimuli until its completion. It will be recalled that each series consisted of 16 separate judgments and that there were eight such series per subject.

$I U_1$  for each subject is the mean number of "equal" judgments per series based on the eight series.

In order to estimate the reliability of  $T_1$ , the correlation with the Time Score from the Density Judgment Problem was determined and will be reported later in this section of this chapter.

The reliability of  $I U_1$  was estimated by determining the split-half reliability coefficient. The scattergram, showing  $I U_1$  for the first 4 series vs.  $I U_1$  for the second 4 series is presented in Figure 1.

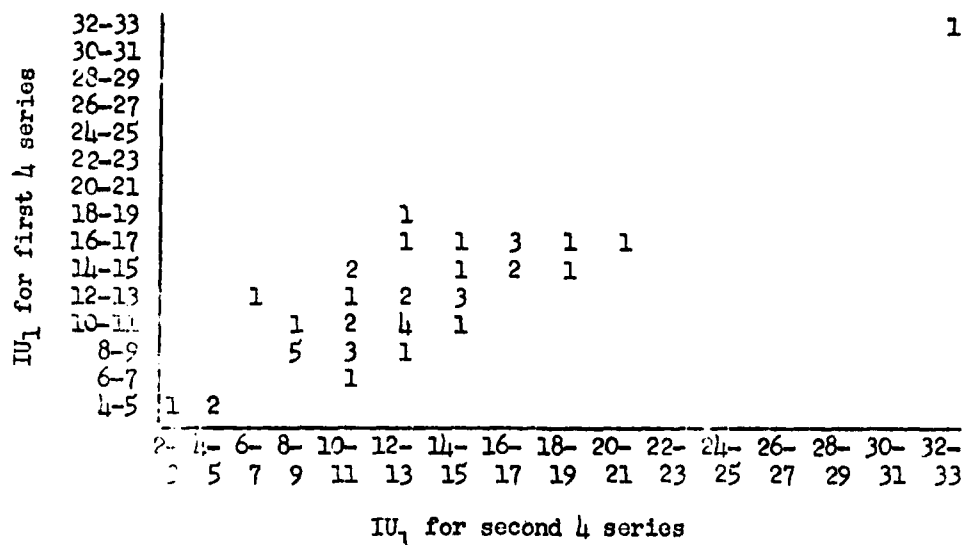


Figure 1. Scattergram of IU<sub>1</sub> for first vs. second 4 series of judgments.

As can be seen, the relationship is large and positive. The uncorrected reliability coefficient is .87, which yields a coefficient of .93 when corrected. In subsequent experimental work fewer series might be used to save time, so reliability of the IU Score should probably be considered to be somewhere between .87 and .93, assuming the use of 4 or more, but less than 8 series. It is probable that the number of judgments per series could be reduced to 10 or 12 without changing the reliability materially, thus effecting a substantial economy of time.

Judgments of Density: Three scores are used here: (a) T<sub>2</sub>, which is defined in the same manner as was T<sub>1</sub>, (b) IU<sub>2</sub>, which is a slightly modified version of IU<sub>1</sub>, and (c) R, which is the mean number of reversals per series, i.e., the mean number of times the subject reversed the established trend within a particular series.

Reversals are scored in a series in which the physical stimuli are increasing when the subject makes a judgment of "loss" after making a judgment of "equal" or "more," or makes a judgment of "equal" after making

a judgment of "more." Conversely, in a series where the physical stimuli are decreasing, reversals are scored when a judgment of "more" is made after a judgment of "equal" or "less" or a judgment of "equal" is made after a judgment of "less."

The split-half reliability of R was found to be .83 uncorrected and .90 corrected. The scattergram is presented in Figure 2.

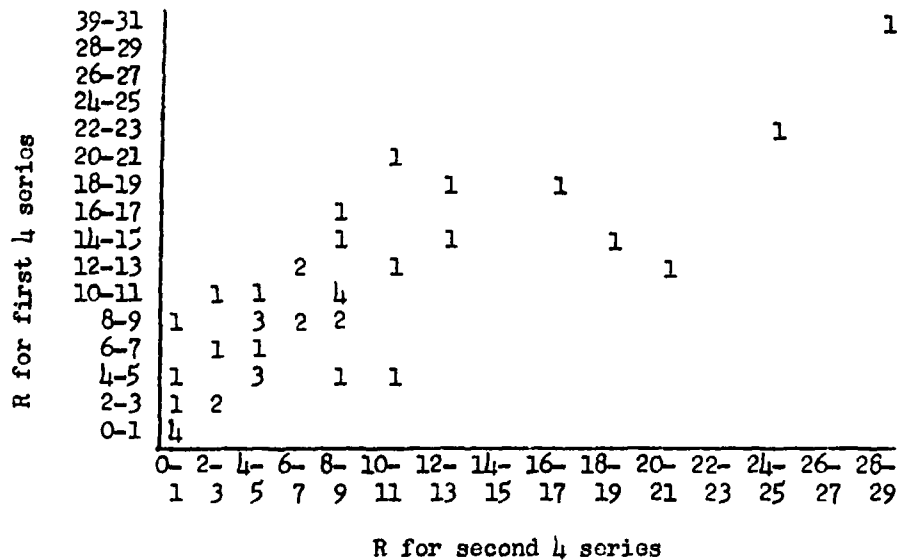


Figure 2. Scattergram of R for first vs. second 4 series.

Between Situations Comparisons: The correlation scattergram of  $T_1$  vs.  $T_2$  is presented in Figure 3. It will be noted that the relationship is substantial and positive.

The Pearsonian product-moment correlation is .74. This can be viewed as the minimum estimate of the repeat reliability of this "Time of Judgment" measure. The experimental procedures were not identical (ease of discrimination between the physical stimuli and sense modality being the major known differences); therefore, this is probably a lower correlation than would be obtained if the procedure were repeated more exactly, at different times.

In order to make this comparison, it was necessary to establish a definition of the Interval of Uncertainty where reversals in judgment are involved. The definition adopted here is that where reversals occur the size of the Interval of Uncertainty for each run is judged on the basis of the number of equal judgments occurring in uninterrupted series

except where only one non-equal judgment occurs in an otherwise uninterrupted series of equal judgments. In this case the size of the Interval of Uncertainty for such runs is taken as the number of judgments in the equal series, including the single interruption. It will be recalled that IU is the average of the number of equal judgments for the runs being counted.

The correlation coefficient between  $IU_1$  and  $IU_2$  is .68, which is significantly different from zero. This can be taken as the estimated minimum reliability coefficient for this measure, and there can be little doubt but that more comparable conditions at different times would yield a higher reliability coefficient, which could be as high as .93, as has been demonstrated in the consideration of the split-half reliability coefficient for IU.

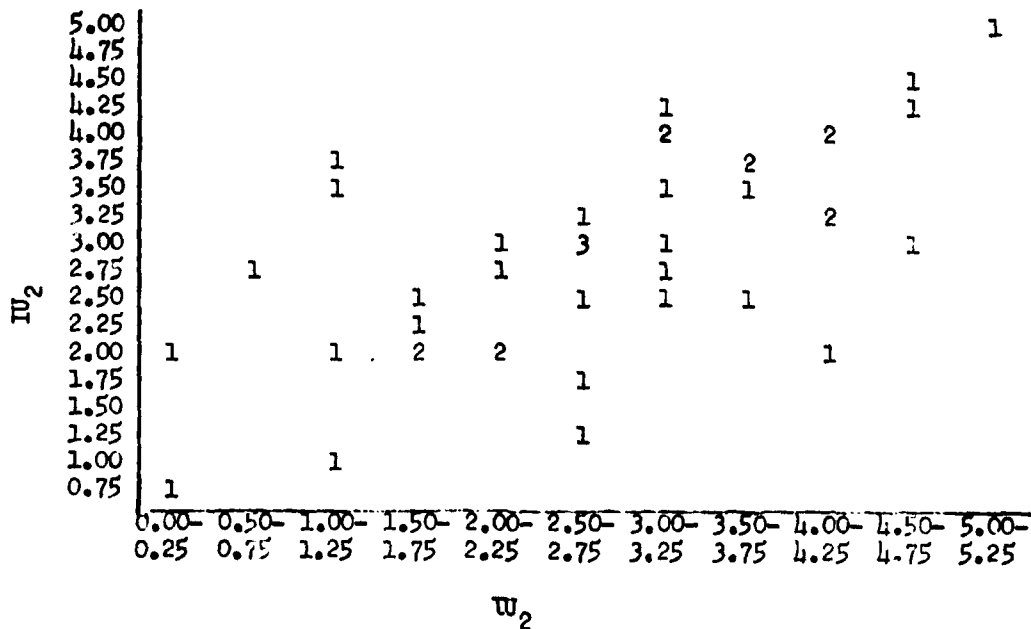


Figure 4. Scattergram for  $IU_1$  and  $IU_2$ .

In the matching of densities problem, the R and  $IU_2$  measures might be correlated, since some number of "equal" responses seems likely if any reversals occur because of the instructions which imply that all three response categories are to be used. There is no correlation between R and  $IU_2$ . In order to investigate the question of the relationship between R and IU thoroughly, it is necessary to determine if any correlation exists between R and  $IU_1$ . These measures appear possibly to be related to a small extent, the correlation coefficient being .21,  $C. R. = 1.37$ ,  $p = .09$ .

Since time scores and scores based on the number of equal judgments per series or number of reversals per series might be correlated and hence might not be independent measures, the correlations between  $T_1$  and  $IU_1$ ,  $T_2$  and  $IU_2$ , and  $T_2$  and R were calculated. The correlations are:  $T_1$  vs.  $IU_1$ , zero order;  $T_2$  vs.  $IU_2$ , zero order; and  $T_2$  vs. R, .47,  $p = .03$  (one-tail test).

#### Discussion

No traditional psychophysical analysis of results has been included here. The reason is that this is not the purpose of this study. The effort in this discussion cannot be to discuss the results in an exhaustive manner, but will be limited to suggesting possible implications and interpretations of the results obtained.

This study was done to determine whether or not the intelligent young adult has a characteristic manner of approach in inconsequential judgment problems. The assumption is that if such a characteristic approach exists, then any other judgment behavior can reasonably be thought of as being made with this fundamental approach being the predisposition operative and decisions are probably to be understood in part in terms of the affect of this predisposition. A first step in establishing the

tenability of this assumption has been taken.

Little can be said, on the basis of these results, as to the generality of these basic judgment behavior patterns. More investigation would be necessary before this point could be clarified, as would also be necessary to determine the extent and manner to which these basic patterns are related to more complex and more structured judgment situations.

The conclusion on the basis of the correlation of  $T_1$  and  $T_2$  must be that to some extent the human being can be viewed as having a "natural" tempo of judgment in these situations. Questions are, as always, immediately raised by this finding. What is the intraorganism generality of "judgment tempo?" How does this type of personal tempo fit into the larger content of other types of personal tempo which have been identified?

Likewise, it is clear that the size of the range of subjective equality is an individual characteristic. The generalizations suggested are that people manifest stable differences in their propensity for utilizing categories of different sizes, or, alternatively, that this is a measure of degree of tolerance for ambiguity, or that people vary in their predispositions to experience things as being similar or different. Here, again, the finding leads to many questions for further investigation.

The phenomenon of reversals is truly tantalizing. The subjects are intelligent students accustomed to grasping quickly complex information. The instructions that reversals were not appropriate were clear and simple. Nevertheless, all except four of the 43 subjects made reversals. A variety of hunches suggest themselves. Maybe the situation was so frustrating as to cause the subjects to become negativistic or possibly fatigue or boredom was great enough that the subjects did not remain alert, but rest intervals were frequent, the difficulty of the task was spoken of by the experimenter so as to allow the subject to verbalize his tensions about it,

and almost all of the subjects gave the impression of trying hard. It is possible that what is reflected is differences in propensity for "changing one's mind," for revision of judgments which are made only with a feeling of difficulty and tentativeness. It is also possible that making reversals reflects a tendency to experience things as similar when doubt exists or a comparison is difficult. The person who makes frequent reversals seems unable to maintain his set as well as can a person who makes few reversals. An alternative way of formulating this explanation is to say that the person who makes frequent reversals is more subject to oscillation about a threshold than is one who makes few reversals. What is being suggested here is that there are significant individual differences in the tendency to oscillation where a response or discrimination is at the threshold, and that the capacity to maintain a standard or set is a reflection of control of oscillation tendencies or a lack of them. A further implication might be that individuals who tend to much oscillation might prefer to attack a task involving various functions by doing a little of this and a little of that, i.e. by alternating sub-tasks, whereas those who tend to less oscillation would manifest little alternation of sub-tasks.

The fact seems to be that reversals do occur and that this is a fairly stable individual tendency. The meaning of this result can only be determined by further investigation.

The possibility that R and IU are to a small extent measures of the same thing(s) is indicated by the correlation found between R and IU. However, since no correlation was found between R and  $IU_2$ , and in view of the significance level of the correlation coefficient between R and  $IU_1$ , it is apparent that R and IU are predominantly measures of different characteristics.

Correlations of zero order exist between  $IU_1$  and  $T_1$  and between  $IU_2$  and  $T_2$ , so the conclusion that these are measures of different psychological functions which are independent seems justified. The significant correlation between  $T_2$  and  $R$  is of moderate size (.47) and indicates that the time measure  $T_2$  is changed by the occurrence of reversals. This indicates clearly that the correlation between  $T_1$  and  $T_2$  was made lower by the occurrence of reversals. The occurrence of reversals may induce stress in part by causing the subject to depart from his natural tempo. The subjects did find the density problem more stressful than the brightness problem.

All of the measures discussed in this section have sufficiently high reliabilities to justify their continued use. While reliability of a measure does not in itself indicate that any important psychological significance is attached, a phenomenon which has little reliability is of little use. Since these measures seem to reflect functions and processes of psychological importance and have reasonably high reliabilities, they can furnish the focus for additional investigation. Some avenues of further exploration have been suggested.

Further study of the intraindividual generality of these phenomena is of importance.

### Chapter III

#### The Concept Evaluation Test

The problems described in Chapter II were undertaken primarily because they offered a means of studying judgment behavior, one aspect of some decision making. We turn now to a method which is conventionally subsumed under the rubric of "personality measures."

In this, as in subsequent chapters, it will become clear that many aspects of personality measurement can be viewed as involving decision making functions as here defined. In undertaking the study of human behavior by concentrating on decision making we have been aware that we will be looking at many partially explored techniques and topics in a somewhat new light. Personality measures are usually interpreted either in terms of personality types or in terms of defenses and complexes. Although the full topic can not be discussed here, we propose to analyze these measures not only in these more traditional terms but also, on occasion, in terms of decision making functions involved.

#### Method

This basic technique was developed by McReynolds (5) and is used here in a slightly modified form. Briefly, the ten ink blot cards of the Rorschach test were presented to each subject(S) one at a time and S was asked to say what he was reminded of or what might be represented by each ink blot. After S had seen and responded to each card he was again shown each one, being asked this time whether or not a specified whole or part of the blot could represent some particular thing or concept. In all, S was asked to say "Yes" or "No" to five concepts per card for a total of 50 concepts.

The concepts were originally selected on the basis of normative data so that a few are seen by almost nobody, a few are seen by almost everybody,

with the rest distributed between on the basis of various proportions of a normal population accepting or rejecting the concept.

The three scores derived from this test will be discussed briefly. The S can respond with 0 to 50 "yes" answers or with as many "no" answers. The number of "yes" responses is represented by a standard score called "J." A low J is the result of few "yes" answers and a high J represents many "yes" answers. Evidence exists to justify the view that low J is associated with strict standards of evaluation in thinking, with highly channelized thought patterns, with the tendency to be ruminative, and with taking things seriously. The converse characteristics have been thought to be associated with high J.

The second score ("V") reflects the consistency with which the S applies his own similarity standards to the various items in the test in terms of the item criteria based upon a normal sample. If V is high the indication is that such standards are applied consistently, while a low V indicates an inconsistent application or oscillation. The V score correlates rather highly with the F+ Rorschach score devised by Beck (1).

On the basis of rather subtle (possibly tenuous) reasoning, a third score, E, is derived on the basis of the difference between V and J. A low E score is thought to be associated with weak ego-control, or what might be called inadequate tension binding, while a high E score is associated with much ego-control.

Moderate (T scores around 50) J, V, and E scores are most frequent. There is some clinical and normative evidence available to support the general ideas presented here.

The test was administered, along with other procedures, at one of the two sessions attended by each subject in this study. There is no reason to believe that the results on this test were affected in any systematic way by the experimental context.

## Results and Discussion

Only those results relevant to the Psychophysical Tasks (see Chapter II) will be reported here.

A number of hypotheses were formulated regarding expected relationships between the measures derived from the Psychophysical Tasks and the measures from the Concept Evaluation Test discussed above in this chapter. These will be discussed in connection with the relevant results.

The existence of a positive relationship between IU (the interval of uncertainty) and J (tendency to answer "yes", i.e., accept the proposed concept) was predicted. IU can be thought of in part as a measure of a tendency to impose structure, i.e., to categorize or label, to lay down rules, etc. More specifically, a low IU would indicate a tendency to be rigid in these regards; the converse being held also. This view of the possible significance of IU would seem to be indicated by the psychological situation. S is given the opportunity to define the equal category either narrowly or broadly. Since there are few, if any, explicit cues guiding S as to the appropriate size of the equal category, and since he has some predisposition to the use of his characteristic size of equal category. (We can not consider here the question of how this predisposition is established.) Of course, it could only be presumed before the study that the IU score would be reliable. But, given this information (see Chapter II) the above view of the significance of IU seems reasonable, if not high in precision. J seems to be a measure of a tendency to apply standards either precisely (a low J) or loosely (high J). Thus, the same general significance is attributable to both J and IU on the basis of a reasoned consideration of the situations. So, the hypothesis was held that a positive relationship would be found between these variables.

The IU score used in this analysis will be that derived from the judgment of brightness situation (see Chapter II), since it is uncomplicated by the occurrence of reversals.

The correlation coefficient for  $IU_1$  vs. J is .30,  $p=.03$  (one tail test). Thus, the degree of relationship is not large but is positive. The hypothesis is supported, but the measures have only a small amount of common variance.

These same results can be evaluated in a different way which is more comparable to designs frequently used in current research as reported in professional journals. This design is as follows. It was reasoned that there would be a positive relationship between IU and J. Those subjects scoring below the overall group mean on J are assigned to a "No" group and those individuals scoring above the overall mean are assigned to a "Yes" group. The IU mean of the "No" group is 2.75 and that for the "Yes" group is 3.37, the difference being significant at the .05 level (one tail test). This difference can be somewhat enhanced by discarding subjects with middle range J scores, assigning individuals scoring high and low to "Yes" and "No" groups and testing the significance of the difference between the IU means for these groups. These results are comparable to those frequently reported and used as the basis for considerable theorizing. A correlational analysis is less spectacular, but probably a more adequate method in such a situation.

There was no expectation of a relationship between V and IU and none of significance was found. It will be recalled that V seems to be a measure of the capacity to apply personal standards consistently. Since IU was considered to be related to a kind of internal standard rather than to the consistency of application, there was no grounds for expecting these measures to be correlated.

A negative relationship was predicted between E and IU. E seems, on the basis of experience with the test, to be related to the degree of ego-control. The usual population shows an Interval of Uncertainty (IU) greater than one judgment interval in situations such as that used in this study. The S who has a low IU is presumably rather rigid and categorical, which is a kind of strong ego-control; the converse is also held. The relevant results tend to confirm the hypothesis in that the correlation coefficient between E and IU is  $-.27$ ,  $p=.04$  (one tail test).

To summarize the relationship between IU and the measures from the Concept Evaluation Test, it has been found that the tendency to equate a larger number of different brightness light stimuli and the brightness of a single tone is associated with the tendency to equate a suggested concept with an ink blot area. The measures presumably reflecting the tendency to be rigid and categorical are associated in the expected way, in small degree.

The relationships of J, V, and E with R (reversals) will be discussed next. No relationship between J and R was anticipated and none of significance was found. It was predicted that Ss scoring low on V (variable in applying subjective standards as measured by the Concept Evaluation Test) would score relatively high on R. The result on this point involves an extremely small number of Ss, but is in agreement with the expectation. The eight Ss who scored lowest on V are all at or above the mean on R. This result can not be evaluated statistically. It was predicted that Ss scoring high on E (over control Ss) would be low on R, but this prediction was not substantiated by the data.

No prediction of the following relationship was made, but the question arose as to whether or not E was related to the time taken in making judgments in the Psychophysical Task. It was found that extreme time scores,

either fast or slow, are associated with extreme E scores, either high or low; while middle range scores on time and on E are associated. This trend is significant at the 2% level of confidence. This finding makes psychological sense, in that a high E score is presumably a measure of rigidity and a rigid person might either take a long time judging to be sure to make the best possible performance, or might form his standard early and so be able to make quick judgments. Special study would be necessary to clarify this situation. Similar factors seem to be involved in the tendency of nonrigid individuals to be either fast or slow in making judgments. These people might make quick judgments because they don't mind making some errors, and others have a firm standard and must reestablish their standard constantly and this takes time. Of course, other hypotheses are possible, e.g., the rigid, fast subject may have developed self instructions that speed is primary and the slow, rigid person may have adopted self instructions emphasizing accuracy.

In overall summary of the results discussed in this chapter, two facts seem clear: (a) there are some consistencies between Psychophysical Task judgment scores and analogous scores based on judgments of the appropriateness of interpretations of ink-blot areas; and (b) these relationships are relatively small in magnitude.

## Chapter IV

### Level of Aspiration

The Level of Aspiration (LA) experiment is usually thought of as a method for studying goal setting behavior, and response to success and failure. Goal setting may be conceived of as a kind of decision and LA as a measure of a kind of decision making. People frequently plan to accomplish certain goals, and then reevaluate these aspirations in the face of later experience. The LA experiment is a model of this type of experience. When the goal is set, common sense usage would dictate our saying that the person is deciding what goal to set and later is deciding if the goal should be revised. LA is, however, not only a situation where goals are set, but has many other facets.

The decision behavior in the LA experiment seems to be quite different, psychologically, from those discussed in the previous chapters except in one way, which will be mentioned below. However, it is seldom safe in psychological investigations to assume anything. Therefore, the LA experiment was included in this series in order to check empirically on its apparent independence from the other types of measures used.

### Method

The same Ss participated in this situation as participated in all other situations reported in this technical report. A typical Level of Aspiration (LA) experiment was conducted, employing a pursuit-rotor as the task. (For a general discussion of the LA experiment see 4.) Briefly, the subject is asked to keep a pointer on a brass target which is embedded in a revolving phonograph disk. The time the pointer is on the disk is recorded electrically and constitutes the score.

After each trial, S is told his score on that trial and is asked to say what he believes he will make on the next trial. The discrepancy between

this last actual score and stated prediction is the measure of the aspiration. These aspiration discrepancies are summed or averaged over all trials, fifteen in this case, and this is the D score. In the study, the median of the aspiration discrepancies was taken as the D score. A score based on the range of individual trial aspiration discrepancy scores was also calculated. This was defined as the range between  $Q_3$  and  $Q_1$  (the third and first quartiles) for each individual. An unfortunate oversight resulted in the time to state the aspiration not being recorded.

#### Results

No relationships of significant magnitude were anticipated between IU, or R scores from the Psychophysical Tasks experiment and the measures defined above for IA. A relationship had been anticipated between T scores, but this could not be tested due to the oversight mentioned above. Those two experimental situations, both of which can be viewed as yielding measures of decision making (or choice or judgment), broadly define two types of decision problem.

One relationship anticipated was between the E score from the Concept Evaluation Test and the D score from IA. It will be recalled that E seems to reflect, at least in part, the appropriateness of the level of ego-control. A high E reflecting over-control and a low E reflecting under-control. It was thought that either extreme of control would tend to result in an extreme D score. It was not possible to predict a simple correlation, however, because, for example, a high E person might either correct his aspiration for every achievement, or he might select an aspiration level and keep it almost regardless of what happened in achievement. In the first case, one would expect a low D, but in the second case the D might be either high or low, depending on other factors. The hypothesis was reached that extreme E scores should be associated with extreme D scores.

The results support the hypothesis stated above. As with other results reported, the relationship is small, but statistically significant (5% level of confidence) and psychologically meaningful. Good balance in judgments of the appropriateness of matching a particular concept with an ink-blot area tends to be present in the same individuals who decide on moderate aspiration levels and modify these levels moderately, whereas extreme behaviors in these activities tend also to be associated. These relationships are small in terms of the proportion of the presumed non-error variance of the IA measures. Study of the reliability and generality of the IA measures is under intensive study as a part of this project.

Since IA literature indicates that measures from this experiment are very much subject to set and situational influences, it was decided that some check should be made to permit a judgment regarding the representativeness of the IA results. Relevant literature and experience indicated that the so-called Zeigarnik effect is associated with representative IA results. The Zeigarnik experiment involved presenting Ss with a series of brief tasks, some of which are interrupted and some of which are completed. A proportion or difference score based on the number of completed tasks recalled as contrasted with the number of interrupted tasks recalled is the usual measure derived from this experiment. The expected relationship between IA and Zeigarnik measures was that higher D (IA) would be associated with the opposite Zeigarnik trend. This expected relationship was found ( $p=.03$ ), and is small as was expected. The IA performances in this experiment seem to be representative, in the sense that they are like those usually reported in the literature.

In considering the level of aspiration experiment as a measure of decision making, the potential significance of the amount of "correction" of the D score in the light of knowledge of performance was considered.

If a person made small "corrections," the spread of D scores would be small, and if "corrections" were large, the D score would be large. In the past, this has been thought of primarily as a reaction to success or failure, the amount of "correction" being assumed to be related to the type of reaction. However, it is possible that this is to an important degree a manifestation of a tendency to be variable in responding or to be relatively invariable in responding. In other words, the possibility exists that, for whatever reason, a person has, at any given time, a pre-disposition to manifest a characteristic degree of variability. In part, this is related to the question of the reliability of the D score. However, if a person does have this characteristic disposition to be variable to some particular degree in responding, then measures from different situations should reflect this tendency.

The LA task involves successive responses where the S selects a score on a scale. A comparable situation was sought and the so-called Autokinetic Experiment was selected. In this task, S sits in a dark room and, after an adaptation period, is asked to look at a point of red light for a few seconds at a time. The light appears to move, though physically it does not. The S is asked to say how far the light moves on each trial.

Since Ss vary widely in the average amount of movement seen, a simple measure of variability would not reflect the characteristic under discussion. The measure desired is a measure of variability relative to the individual mean amount of movement seen. Therefore, the coefficient of variability was calculated for a series of 20 judgments for each subject.

The variability score ( $Q_3 - Q_1$ ) from the LA situation and the variability score from the Autokinetic situation were correlated. There is a positive relationship of moderate size between these two scores. The rank-order correlation of these scores was calculated for each examiner's group

of Ss and an average of these correlations was determined (using the Fisher z transformation). The mean correlation is .64, which is significantly different from zero. This method of estimating the size of the correlation coefficient was used because the scattergrams revealed that there were probably examiner influences which could not be removed without making undue assumptions.

It seems clear that in the type of decision making where goals are set and revised repeatedly, there is operative a tendency, characteristic of each individual, to make small, moderate, or large revisions when the objective circumstances are highly similar. This is a problem requiring further study. It would seem fair to conclude, at this point, that decision making of this general type can not be fully understood unless this variation tendency and its relationship to decision making can be further understood.

## Chapter V

### Examiner Influences

The results reported in this chapter were not anticipated. The persons serving as examiners in these situations were graduate students who had completed at least one year of graduate work and who were trained in the techniques used. All instructions and procedures were standard.

Each S was assigned to one examiner who conducted all procedures with his Ss. Four examiners were involved, and 45 Ss. Examiner F processed 14 Ss, examiner H, 12 Ss, examiner J, 9 Ss, and examiner W, 10 Ss. This method of procedure was chosen because it reduced scheduling difficulties, which are inevitably great when Ss are seen more than once, and especially where they must be seen for 5 hours, total, as in this study. It has usually been assumed that trained examiners would not be found to have influenced their Ss' performances differentially. However, in order to be sure that no such differential influences had been operative, appropriate analyses of the data were conducted.

The Ss seen by the four examiners were treated as four separate groups and an analysis of variance of group mean differences was computed for series of variables considered potentially most susceptible to examiner influence. The following results were obtained: (a) the T (time) score from the Psychophysics Tasks is significantly affected by the examiner, (b) the IU score is probably not significantly affected, although the F ratio is close to the .05 level, (c) the R (reversals) score is not affected by examiner influences, (d) the V score is significantly affected by the examiner, (e) the E score is significantly affected by the examiner, (f) the D score of LA is significantly affected by examiner.

The author is satisfied that proper precautions were taken to insure comparable procedures, so he must conclude that results of psychological

experiments, especially where the experimenter must be fairly active with the subject as in procedures reported here, are frequently affected unknowingly by the experimenter in ways usually unknown.

Results reported in previous chapters have been arrived at after corrections for examiner differences had been made. Subsequent work using similar procedures will be conducted having one experimenter see all Ss in a given procedure.

A more complex type of examiner influence is suggested by some of the results. If correlations between variables are calculated separately for each examiner's Ss, these are so different in some instances as to suggest that not only the amount of relationship but even the direction (direct or inverse) of relationship is different for different examiners. Since the examiner groups are small, these trends must be studied on larger samples if conclusive results are to be available.

Assuming such results are substantiated, the implications for objective experimental and measurement techniques in the personality area are rather disturbing. One conclusion would have to be that rather than measuring an S we are frequently measuring a relationship.

## Chapter VI

### The Ego-Control Variable

Before any of the work reported here had been undertaken, preliminary studies had been conducted. The work reported in previous chapters had proceeded only after the preliminary studies had indicated some promise that these procedures would yield pertinent information.

Preliminary studies had also indicated that the ego-control conceptualization of personality was a relatively powerful theoretical point of view in considering decision making. Since the ego-control conceptualization is still in a rudimentary form, study directed toward clarification of this viewpoint was considered to be appropriate. Therefore, at the same time the investigation reported in previous chapters was being carried out, a study of some measures of ego-control was made.

The ego-control variable was most clearly discussed in work by Jack Block (2) and Jeanne Block (3). They had studied a sample comparable to that being used in this project. The methods they developed for determining the type of ego-control are very time consuming. The need was to determine whether simpler methods could be used to measure the type of ego-control. There was no assurance that any simpler methods would or would not be successful.

### Method

In order to use the Block sample as the criterion group in testing simpler procedures, thus saving hundreds of experimental hours, it was necessary to make the contacts by mail because the sample was scattered geographically. The simpler procedures selected were two pencil and paper "personality tests" which had yielded useful results in connection with discriminations somewhat similar to that being attempted. These techniques are the Minnesota Multiphasic Personality Inventory (MMPI) and

the Adjective Check List (ACL). Both are readily self-administered by people of the kind in the sample.

Each S was mailed a letter asking cooperation, the test materials, and a self-addressed, stamped return envelope. As added inducement, a nominal payment was made for the time spent in completing the tests. Approximately 65% of the original sample cooperated.

#### Results

The results for the Adjective Check List were analyzed in two ways. First, the total number of adjectives checked as being descriptive was determined. The mean for the nine under-controllers (UC) who responded was 98 words, the mean for the 13 adequate-controllers (AC) was 71, and the mean for the 16 over-controllers (OC) was 82. An analysis of variance treatment of these results showed that the means are significantly different (5% or greater level of confidence). This is an impressive result, considering the small number of Ss. Tests (t) of the significances of the differences between individual group means were made, it being found that the UC - AC means are clearly different ( $p < .01$ ), and the UC - OC means differ ( $p < .05$ ), and the AC - OC means are possibly significant though not clearly so ( $p > .05$ ). The conclusion reached is that the total number of checks on the Adjective Check List could be used, after appropriate refinements, as a method for determining the type of ego-control.

The second type of analysis for ACL consisted of determining whether or not there were clear clusters of specific adjectives used by most members of each group and not by any members of the other groups. All lists were compared systematically and all checked adjectives used by members of different groups were discarded. A cluster of discriminating, non-overlapping adjectives was found for each of the three ego-control groups. No statistical test was made, but the fact that such lists were found indicated

the strong probability that this type of scoring would further strengthen the power of the Adjective Check List as a method for making the discrimination in question.

The results obtained with the MMPI indicate that it too is a technique which could be adapted for use in making the ego-control differentiations. The nine so-called clinical scales and a "Social Introversion" scale were studied for differences between groups by means of an analysis of variance. Significant differentiation among the groups was found on the Hy, Pa, Pt, Sc, and Ma scales and also on the Social Introversion (Si) scale. The comparative profiles are shown in Figure 5. An item refinement is in process and it is clear that an ego-control scale can be developed.

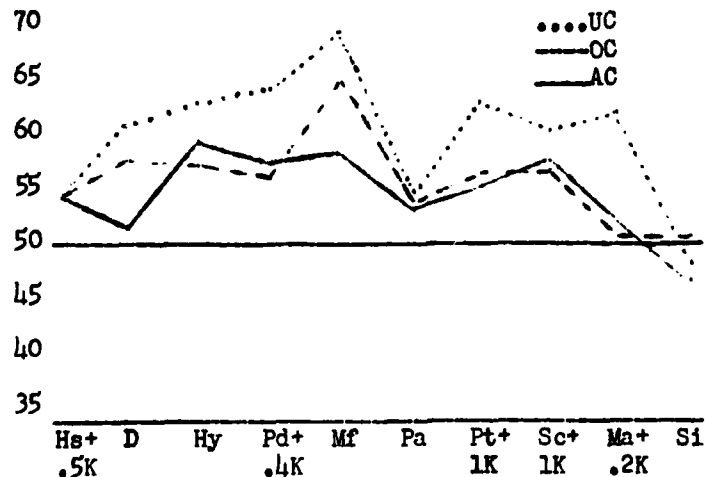


Figure 5. MMPI results for three ego-control types: UC indicates Undercontrol Group, OC indicates Overcontrol Group, and AC indicates Adequate Control Group.

These results indicate that simple inventory type tests can be developed which will permit ready classification of individuals as to their type of ego-control. This development of these techniques should be continued, so as to permit clearer, broader, economical use of the ego-control concept in the analysis of decision making, as well as in other research.

## Chapter VII

### Conclusions

The basic measures of decision making used in this study are from the psychophysical type situations, involving the matching of graded series of stimuli to a standard stimulus of a different modality. It has been shown that for decisions of this type there are characteristic individual differences in time for making decisions, in the size of equal intervals utilized, and in the ability to maintain a subjective standard. Further study will be necessary to establish the intraindividual generality of these characteristics, and to establish the relationships between these characteristics and decision making characteristics in more complex situations.

A variety of so-called personality measures were investigated and many of those which were expected on the basis of reason to be related to the various decision making characteristics were found to be related in the expected way, but typically these relationships are relatively small in magnitude. One might expect the personality measures to be more strongly related to decision making characteristics in more complex situations.

Bibliography

1. Beck, S. J. Rorschach's test: I. Basic processes. New York: Grune and Stratton, 1944. II. A variety of personality pictures. New York: Grune and Stratton, 1945.
2. Block, Jack. Ph. D. dissertation. Stanford University Library.
3. Block, Jeanne. Ph. D. dissertation. Stanford University Library.
4. Lewin, K., Dembo, Tamara, Festinger, L., and Sears, Pauline S. Level of aspiration. In J. McV. Hunt (Ed.), Personality and the behavior disorders. New York: Ronald Press, 1945. Pp. 333-378.
5. McReynolds, P. Perception of Rorschach concepts as related to personality deviations. J. abn. soc. Psychol., 1951, 46, 131-141.